Daily precipitation over southern Africa: a new resource for climate studies.

R. Layberry

University of Sussex, UK.

D. R. Kniveton*

University of Sussex, UK.

M. C. Todd

University College London, UK.

C. Kidd

University of Birmingham, UK.

T. J. Bellerby

University of Hull, U.K.

*Corresponding author address

Department of Geography

University of Sussex

Falmer, East Sussex, BN1 9RH

UK

Email: kafw3@susx.ac.uk

Abstract

This paper describes a new high-resolution multi-platform multi-sensor satellite rainfall product for Southern Africa covering the period 1993-2002. The Microwave Infra-Red Algorithm (MIRA) employed to generate the rainfall estimates combines high spatial and temporal resolution Meteosat infrared data with infrequent Special Sensor Microwave Imager (SSM/I) overpasses. A transfer function relating Meteosat thermal infrared cloud brightness temperatures to SSM/I rainfall estimates is derived using co-located data from the two instruments and then applied to the full coverage of the Meteosat data. An extensive continental scale validation against synoptic station data of both the daily MIRA precipitation product and a normalized geostationary IR-only GOES Precipitation Index (GPI) demonstrates a consistent advantage using the former over the latter, for rain delineation. Potential uses for the resulting high-resolution daily rainfall dataset are discussed.

1. Introduction

The availability of water in southern Africa is spatially highly variable (Houghton *et al.*, 2001). Controlled primarily by rainfall, water resources vary from abundant in the tropical zones in central Africa to scarce in the south west of the sub-continent. Even in countries where water resources are generally relatively abundant, inter-annual variability of rainfall can be considerable. For instance, Mozambique experienced drought conditions in 1998 and severe flooding in 2000 and 2001. While the importance of information on precipitation is not in doubt, much of region suffers from inadequate measurements. In Figure 1 the spatial distribution of daily reports of rainfall from the GTS network for the period 1990 to 2000 are shown. The figure shows large areas over much of southern Africa where there are little or no measurements of daily rainfall, most notably over Angola and the Democratic Republic of Congo. It is within this data void that satellites can provide vital information on precipitation. The following work is applied to the area of southern Africa indicated on Figure 1 bounded by longitudes 10° and 50° East and latitudes 0° and 35° South.

The science of satellite precipitation retrievals was first established over a quarter of a century ago using data from the infra-red (IR) (10.6-12.6μm) and visible (VIS) (0.4-0.7μm) portions of the electromagnetic spectrum. Techniques using these data are based on the assumption that clouds with high tops (discernable from low IR temperature brightness), and substantial vertical depth (bright in the visible wavebands) are most likely to precipitate. This assumption is most effective for convective conditions, where the majority of the rainfall comes from tall, strongly precipitating cumulonimbus. However, even in strongly convective regimes complications remain due to the presence of high, non-precipitating cirrus. Methods based on thermal IR imagery alone rely on empirical relationships derived between cloud characteristics (e.g. cloud top temperature) and surface rainfall (for a review see Kidd*,* 2001). The GOES Precipitation Index (GPI) (Arkin and Meisner, 1983) is perhaps the most widely used example of such 'cloud-indexing' methods. The launch of the Special Sensor Microwave Imager (SSM/I) in 1987, on board the Defense Meteorological Satellite Programs (DMSP) 5D-2 spacecraft F-8 increased interest in satellite based precipitation retrievals. Unlike techniques based on VIS/IR measurements, passive microwave (PM) data from SSM/I allowed a physically more direct means of monitoring rainfall due to the attenuation of upwelling radiation by hydrometeors themselves and precipitation related ice particles. The physically more direct nature of the relationship between satellite PM measurements and rainfall was extended further with the launch of the Tropical Rainfall Measuring Mission (TRMM), in 1997, with a precipitation radar (PR) instrument on board. A number of international intercomparison projects have attempted to assess the degree of accuracy possible with satellite data based precipitation algorithms (Barrett *et al.,* 1994; Ebert *et al.,* 1996; Smith *et al.,* 1998; Adler *et al.,* 2001). These projects have shown that PM estimates produced the best instantaneous results.

Unfortunately, although PM sensors are able to provide accurate estimation of instantaneous rain rates, they are mounted on low earth orbiting satellites, which provide poor temporal sampling. This means that PM data-based techniques are most suitable for estimation of accumulated rainfall over longer periods of perhaps a month or more. By contrast, IR imagery from geostationary satellite systems has a higher temporal resolution, resulting in a reduction of the sampling errors at all temporal scales (New *et* *al.*, 2000). To account for the limitations inherent in both the PM and IR precipitation estimates combined IR-PM techniques have been developed (Adler *et al.*, 1993; Huffman *et al.*, 1997; Xu *et al.*, 1999; Bellerby *et al.*, 2000; Sorooshian *et al.,* 2000; Millar *et al.,* 2001; Todd *et al.*, 2001; Joyce *et al.,* 2004). International intercomparison studies have illustrated that combined IR-PM techniques are capable of providing high spatial resolution rainfall estimates at daily timescales with greater accuracy than the IR only methods (Ebert *et al.,* 1996; Adler *et al.,* 2001). In this paper we introduce a 10-year daily rainfall dataset for southern Africa on a 0.1 degree grid, produced using one of these combined algorithms, the Microwave-Infrared Rainfall Algorithm (MIRA) (Todd *et al.*, 2001). Although Todd *et al.*, (2001) provide results of an extensive validation of MIRA over a range of space/time scales, the validation at daily timescales was restricted to a rather limited region covered by the EPSAT gauge network (Lebel and Amani, 1999). Here, we analyze the performance of the MIRA over the entire subcontinent of southern Africa, and we are able to describe some of the characteristics of daily rainfall variability on a fine grid over the region.

2. Methodology

a) Satellite data for the MIRA product

Infrared data from Meteosat and passive microwave derived rainfall data from SSM/I were used in the construction of the daily rainfall rates over southern Africa between longitudes 10° and 50° East and latitudes 0° and 35° South for the years 1993-2002 at a resolution of 0.1°. The Meteosat high resolution (5km) IR data used were at 2 hourly intervals and were obtained from the EUMETSAT Archive Retrieval Facility for the years 1993-1995. The data were converted from raw counts to brightness temperatures (T_b) and projected onto a latitude/longitude grid at 0.05 degrees. The re-gridding scheme consisted of converting the Meteosat coordinates of each pixel into latitudes and longitudes and calculating the mean T_b of those pixel values falling within the grid box of a particular latitude and longitude. Some corrupt data was identified and removed. Additionally, Meteosat high resolution (5km) IR data at 2 hourly intervals were obtained from the TAMSAT group at the University of Reading for the years 1996-2002. This data had already been converted to T_b on a latitude/longitude grid and quality controlled.

Instantaneous rainfall estimates were obtained from SSM/I data using the Goddard Profiling Algorithm (GPROF) (Kummerow and Giglio, 1994; Kummerow *et al.,* 1996, Kummerow *et al.,* 2001). The GPROF algorithm is an inversion type algorithm providing estimates of instantaneous rainfall rates, the vertical structure of precipitation and the associated latent heating. It achieves this by first constructing large databases of cloud model derived profiles, then producing radiative transfer calculations at cloud model resolution. Sensor resolution average quantities are determined by convolving the high resolution T_b field to the observed resolution using antenna gain functions. Using a Bayesian inversion method the algorithm produces a weighted sum of profiles whose T_b signatures are similar to those observed (Adler *et al.,* 2003). The time period of interest (1993-2002) was covered by the F10 and F14 satellites which were subsequently inter-calibrated by the comparison of simultaneous readings from the F10 and F14 and the coincident overlap of the F11, F13 and F15 satellites. The data were obtained at a resolution of 0.5°. In addition to instantaneous passive microwave based rainfall data, a monthly diurnally corrected SSM/I rainfall product was used to normalise the daily rainfall data (Andersson *et al.,* 2003). The diurnally corrected dataset was derived with the aid of data from the Tropical Rainfall Monitoring Mission. Unfortunately TRMM data are only available from 1997. There are two sensors on board TRMM that provide information on rainfall; a passive microwave radiometer of the same type as the SSM/I, known as the TRMM Microwave Imager (TMI), and an active microwave sensor, the precipitation radar. For each of these instruments there are operational algorithms, which provide estimates of rainfall. The TMI rainfall is estimated using the Goddard Profiling Algorithm (Kummerow *et al.,* 2001). In addition, a rainfall product based on a combination of PR and TMI is available, where the PR algorithm is optimized for the distribution of rainfall particle sizes given by TMI (Haddad *et al.,* 1997). However, the PR has a rather narrow swath (220km) such that the sampling in time is very limited. The TRMM satellite is low-earth orbiting, with a non sun-synchronous orbit such that every part of the diurnal cycle is sampled for each location on the Earth's tropical surface over the course of 23 days at the equator and 46 days at the highest latitudes (38°N and 38°S). Rainfall estimates from TRMM if averaged over sufficient time are therefore free from systematic sampling error associated with the diurnal cycle of rainfall. Removal of diurnal bias associated with the SSM/I based estimates in the monthly dataset was achieved by calculating the ratio of the average rainfall for the region from the SSM/I estimates (derived only at SSM/I overpass times) to the average daily rainfall calculate from TMI for each month and removing this from the SSM/I based estimates.

In addition to ensure the diurnally corrected SSM/I monthly rainfall estimates have zero bias with reference to a benchmark, co-temporal and collocated estimates of rainfall from TRMM PR and SSM/I were compared. The mean bias was derived and removed.

b) The MIRA algorithm

The following describes the step by step process used to construct the MIRA daily rainfall dataset.

1) For every 0.5° by 0.5° grid cell over the study region , for each month from 1993 - 2002, the cloud top T_b from Meteosat and the PM instantaneous rain rates from SSM/I were binned for samples where the Meteosat T_b and PM rain rate data were observed within 30 minutes of each other. This gives a large sample of T_bs and associated rain rates within each grid cell from which to derive a T_b to rain rate transfer function, although a significant amount of lower resolution PM data is not used due to the 30 minute threshold for acceptance. The PM rain rate to T_b transfer function is calculated using a method known as histogram matching and described below.

2) For each grid cell the histogram of both T_b and rain rate for an area of 2.5° by 2.5° centred on that grid box was derived. In some cases, the number of points in the rain rate histogram was insufficient to build a representative histogram (<200), in which case the 2.5° grid box was allowed to expand symmetrically in steps of 0.5° in each direction until sufficient points were obtained. This was rare except in very dry areas in the drier seasons where the area would expand until it encountered an area of higher rainfall. While the choice of the exact number of points used to construct the histogram is arbitrary we found too few values gave a

stepped function, too many and the box had to expand to find the required amount of values, meaning that the relationship is gathered over a larger area

3) The histograms of T_b and rain rate were converted to cumulative histograms by integration. Specifically, the histogram of T_b (number of observations of each T_b plotted against T_b) was converted to the proportion of data points which exist below a certain T_b plotted against T_b. Similarly the histogram of rain rate (number of observations of each rain rate plotted against rain rate) was converted to the proportion of data points, which exist above a certain rain rate plotted against rain rate. It should be noted that in coastal locations T_bs over land and T_bs over ocean are included in the same histogram with the assumption that the relationship between rain rate and T_b is the same for both surface types.

4) The histogram matching method was applied, whereby, the T_b associated with each rain rate is the T_b at which the cumulative histogram of T_b is equal to the cumulative histogram of the rain rate. For example, where the value of each histogram is 0.5, the T_b and the rain rate can be read off and associated with each other. Over all values, this gives the transfer function f, where rain rate = $f(T_b)$ for each 0.5° grid box for each month. Figure 2 shows an example of a T_b - rain rate relationship.

5) The spatially (0.5°) and temporally (monthly) variable function f was then applied to the Meteosat IR T_b data at full resolution (2 hourly and 5 km) for the full region

(10º to 50º E and 0º to 35º S and 1993-2002). The final rain rates were averaged over each day, binned to 0.1º by simply averaging of 0.05° grid box values and normalized such that the mean monthly rainfall estimates over the entire study area were equal to the mean monthly rainfall estimates from the diurnally corrected SSM/I dataset described above. The resulting dataset is referred to as the MIRA rainfall estimate dataset.

6) An additional dataset of precipitation estimates using the GPI was created for comparison. The dataset was constructed by applying the simple rainfall algorithm to the Meteosat IR T_b data at full resolution [if $T_b > 235$ K then rain rate = 0 and if $T_b <= 235$ K then rain rate = 3mmhr^{-1}]. Again, the final rain rates were averaged over each day, binned to 0.1º and normalized such that the mean monthly rainfall estimates over the entire study area were equal to the mean monthly rainfall estimates from the diurnally corrected SSM/I dataset. The resulting dataset is referred to hereafter as the normalised GPI.

The sampling resolution for the MIRA product (10km) is finer than its effective cell size (0.50 degrees). The product was generated at a high spatial resolution in order to provide the user with maximum flexibility. For example, rainfall estimates may be aggregated to yield mean areal precipitation within a set of river basins or sub-basins. Of course, such an aggregation process will reduce the variability of the resulting precipitation product to some extent. However, this effect will be offset by the spatial correlations present between neighboring 10km estimates. Validation statistics presented in this paper are for a 0.5-degree spatial resolution aggregated product.

c) Validation data and methods

Validation of MIRA estimates at sub-continental scales requires a spatially extensive set of independent data at daily timescales. The most appropriate source of such data is the Global Telecommunication System (GTS) rain gauge dataset. This dataset contains daily rainfalls interpolated to 0.5º for the African continent. Each 0.5º by 0.5º grid box contains the interpolated daily rainfall total and the number of gauges contained within that grid box. The gauge density is greatest in South Africa and variable elsewhere, with some large areas exhibiting very limited gauge coverage, notably, Angola, Democratic Republic of Congo and Mozambique. This can introduce serious error when interpolating into a significant void using gauges in different climate regimes. In this study, therefore, only data grid boxes with non-zero numbers of gauges were used. The proportion of grid cells with one or more gauges within the area of interest was 5 %, with only 0.5 % having more than 1 gauge.

For comparison, the MIRA and normalised GPI estimates were smoothed and resampled to 0.5º. For each day, the coincident grid boxes of MIRA, normalised GPI and GTS (where non-zero numbers of gauges existed) were collated and comparisons made between MIRA/GTS and normalised GPI/GTS. The number of coincident points for analysis per day was of the order 200-300. Firstly, a contingency table was constructed and a statistical analysis performed for each year. The contingency table compares estimated (MIRA, normalised GPI) and observed (GTS) rainfall in the following ways.

For some rainfall threshold (0.01 $mmhr^{-1}$) each point is either estimated to rain or not and is either observed to rain or not. This gives four outcomes: estimated rain/ observed rain; estimated rain / observed no rain; estimated no rain / observed rain; estimated no rain / observed no rain. These are referred to respectively as hits (h), false alarms (f), misses (m) and zero zeros (z). Various scores assessing the skill of the rainfall algorithm to identify rain can then be derived from these. The following measures are popularly used: Accuracy; Bias; Probability of Detection (POD); False Alarm Ratio (FAR); Critical Success Index (CSI); Equitable Threat Score (ETS); Hansen and Kuipers Discriminant (HK); Heidke Skill Score (HSS); Odds Ratio (OdR) (Stanski 1989). The following are the equations used in the analysis.

Accuracy = (h+z)/(h+f+m+z)

BIASscore = (h+f)/(h+m)

POD = (h)/(h+m)

FAR = (f)/(h+f)

CSI = (h)/(h+f+m)

ETS = (h-expected_correct)/(h+m+f-expected_correct)

Where expected_correct = (f+h)*(m+h)/(z+f+m+h)

HK = (h)/(h+m)-(f)/(f+z)

HSS = 2*(h*z-m*f)/((h+m)*(m+z)+(h+f)*(f+z))

OdR = (h*z)/(m*f)

3. Results

a) IR rain/no-rain threshold values.

During application of the algorithm, the function rain rate = f(T_b) was obtained. Within this function we have information about the threshold T_b i.e. the temperature below which we assume rain occurs. This threshold temperature varies spatially and temporally reflecting the variable relationship between cloud top temperature and surface rainfall, and is in contrast with the fixed value of 235K used in the GPI. This threshold temperature shows a marked seasonal cycle, being higher in the local summer. Over the southern African region as a whole the threshold temperature has an annual mean of 241K and a seasonal range of approximately 20K. Figure 3 shows the mean spatial variation in threshold temperature for December-January-February (DJF) over the 10 year period. There is structure to the pattern of IR thresholds indicating spatially coherent variations in the relationship of cloud top temperature and rainfall and therefore the cloud/rainfall processes. This structure does not appear to be associated with that of the mean rainfall (Figure 5a). There is also considerable interannual variability in the magnitude of IR thresholds in the DJF wet season, although the spatial patterns remain relatively consistent (not shown).

b) Comparison with ground based GTS rain gauge data

Rain gauge data presents the only ground based validation source for satellite based rainfall estimation over the majority of southern Africa. Unfortunately rain gauges are not without error themselves when measuring precipitation due to interactions of the gauge and their micro-environment. Additionally, as mentioned above, gauge data over much of the subcontinent are sparsely distributed. A number of authors have explored the issue of the contribution of sub-sampling by gauges to gauge-satellite differences (Ciach *et al.,* 2003; Gebremichael *et al.,* 2003). In this study we have made no attempt to separate gauge and satellite errors and future research should attempt to deconvolve the contributions to differences between MIRA and gauge representations of the rain field. Part of the error apparent in the MIRA data will arise from the PM data used to define the T_b-rain rate relationship. A large number of PM rainfall algorithms have been developed for use with SSM/I and TMI data with different error characteristics. The GPROF algorithm as applied to the TMI has been shown to overestimate rainfall over land, as shown by a positive bias of 17% when compared to rainfall measures derived from 6700 rain gauges globally, produced by the Global Precipitation Climatology Centre of Deustcher Wetterdienst (Kummerow *et al.,* 2001). However it should be noted that the majority of these rain gauges were located over industrialized countries.

Table 1 presents the overall statistics of MIRA and normalised GPI vs. the GTS dataset, for a typical year 2000. It can be seen from this table that the MIRA method is better than the normalised GPI at identifying raining from non-raining grid cells. It can also be seen (from the value of Bias and OdR) that MIRA tends to over-estimate rainfall area whereas normalised GPI tends to under-estimate. This leads to MIRA having a greater POD and FAR. The Heidke Skill Score (HSS) shows the fraction of correct estimates after eliminating those that would have been correct due purely to random chance. A value of 0 indicates the estimated is random, whereas a value of 1 indicates perfect agreement. Any value greater than 0 therefore indicates the method is ‘skilled’. In this case, the result from MIRA is better than that for normalised GPI, a condition which holds for all years with similar improvement in MIRA compared to normalised

GPI. Figure 4 shows the MIRA-GTS daily POD for 2000. From this figure it can be seen that there is a far better agreement between gauges and satellite estimates in the wetter months than in the drier ones. This is because of the tendency to 'over-predict' (seen in a higher Bias) when it is very dry, leading to a high FAR in these months. Similarly, plots of CSI, ETS, HK and HSS show better agreement in the wet months. The results for normalised GPI-GTS are visually very similar. It should be noted that the results of this comparison are not greatly affected by the rain: no rain threshold chosen with further processing showing little difference between a threshold of 0.01 $mmhr^{-1}$ and 0.1 $mmhr^{-1}$.

Table 2 shows the values of the overall HSS for the 8 years of the survey where the gauge data existed. It can be seen from this table that there is a positive correlation between the HSS skill of both satellite methods and the mean number of grid boxes used in the comparisons (shown in Table 2 and dependent on the number of reporting gauge stations). A higher number of gauges leads to a greater agreement between satellite methods and GTS gauge observations. This is likely due to the higher number of gauges reducing the problems of spatial sampling in the gauge dataset. It also indicates that a proportion of satellite 'errors' in relation to the GTS gauges is associated with poor gauge density in the validation GTS dataset. An alternative explanation for the apparent positive correlation between gauge population numbers and HSS scores may be that the additional grid squares brought into the validation by the increase in gauge population are systematically located in "easier" regions.

Overall, the MIRA algorithm gives a statistically significant (at the 95% confidence level) accurate estimate of rain occurrence, as does normalised GPI. To assess whether MIRA is significantly better at rainfall delineation, than normalised GPI, the HK scores for the two algorithms were compared. By assuming the false alarm and miss rates of the algorithms are independent, the standard error in the HK skill score is the root of the sum of the squared standard errors in the miss and false alarm rates. This leads to a standard error in HK skill score of $\ll 0.01$ due to the large number of 'events' over the course of a year. Thus it can be concluded that the skill score suggests that MIRA is statistically significantly better than normalised GPI at estimating rain occurrence, at above the 95% confidence level.

The ability of MIRA to capture the spatial variation of rainfall can be seen in Figure 5a which shows the mean monthly rainfall over Southern Africa during a representative wet season month (January 1999) derived from MIRA compared with that estimated by the normalised GPI (Figure 5b) . It can be seen that MIRA appears to identify finer detail in the spatial structure of rainfall. Qualitative comparison with the coincident GTS (Figure 5c) data indicates that the spatial structure of the MIRA estimates better represents that of the GTS gauge data than does the normalised GPI. This is perhaps most notable over Eastern SA between 30-35°E and 10-25°S and over coastal eastern South Africa, where gauge density is relatively high. Notably, there is weaker agreement between both MIRA and normalised GPI with GTS in regions where the density of gauges is low (see Figure 1) over Angola and the Democratic Republic of the Congo for example. Figure 6 shows a scatterplot of the MIRA estimates of rainfall for the year 2000 verses those from the GTS gauges at 0.5° for grid squares where there is at least one rain gauge present.

When comparing MIRA and normalised GPI daily rainfall amounts of rainfall at 0.5° resolution, MIRA shows less improvement on normalised GPI. For the year 2000, MIRA has a correlation coefficient of 0.38, a mean absolute error of 0.12 and a root mean squared error of 0.37 compared with the GTS data, while normalized GPI displays a correlation coefficient of 0.23, a mean absolute error of 0.13 and a root mean squared error of 0.34. However in the year 1999, normalized GPI performs better, than MIRA, when compared to the GTS data with a correlation coefficient of 0.26, a mean absolute error of 0.11 and a root mean squared error of 0.27. MIRA statistics for 1999 are a correlation coefficient of 0.22, a mean absolute error of 0.11 and a root mean squared error of 0.34. This suggests that while the MIRA algorithm is better at delineating rain from no rain (as indicated by the skill scores) it does not offer any consistent improvement over normalized GPI in terms of estimating rain amount.

4. Potential Applications of the Dataset

The MIRA algorithm was used to generate daily rainfall maps at 0.1º over southern Africa for the years 1993-2002. These maps have higher spatial and temporal resolution than the SSM/I monthly 0.5º degree maps often used for rainfall analysis over these gauge data sparse areas. Whilst, the rainfall estimated from the MIRA algorithm is by no means perfect, owing to the physically indirect relationship between cloud top temperature and rainfall, the technique dynamically accounts for variations in cloud/rainfall relationships by using a variable calibration scheme, with useful improvements in accuracy relative to the IR-only normalized GPI. The resulting MIRA rainfall product has a number of potential applications, some of which are discussed below.

It is possible with this dataset to record high rainfall events over time periods short enough to be important for studies of localized flooding. Figure 7 shows the average rainfall rate for a high rainfall event captured between 21st and 25th February 2000. This period coincides with hurricane Eline entering Mozambique from the Indian Ocean and combined with high rainfall in the preceding weeks, resulted in widespread flooding and over a million residents becoming homeless in the region. The MIRA integrated rainfall map clearly shows how Mozambique bore the brunt of the disaster and the high spatial resolution allows the integration of rainfall over river catchment sub-basins.

It is also possible with a daily dataset to analyze statistical properties of the data such as the variability of the daily rainfall distribution. Figure 8 shows the coefficient of variation (COV) of the daily rainfall over southern Africa for the entire period 1993-2002. The COV over the Mozambique channel is higher than surrounding areas, possibly associated with the passage of tropical cyclones in this region.

Additionally, hydrological models of large basins require estimates of rainfall at the highest possible spatial/temporal resolution. The MIRA dataset has already been tested in a hydrological modeling application for the Okavango river in western southern Africa (Anderssen *et al.,* 2003). Moreover, hydrological models can be designed to utilize information of the frequency and persistence of rainfall to constrain estimates of evapotranspiration. For example, interception and evapotranspiration losses can be suppressed during rainfall of extended duration. We have derived the probabilities of a rain day followed by a rain day and a rain day followed by a dry day for each grid cell. Figure 9 shows the difference between the probability of a rain-rain day in an El-Nino event minus the same probability in a La-Nina event for the entire period 1993-2002. A definite spatial pattern is evident with a higher probability of a rain day followed by a rain day in an El-Nino year in the north of the region and a higher probability of this in the south of the region for La-Nina years, reflecting the spatial variation of teleconnections with El Nino/La Nina in the region (Camberlin *et al.,* 2001).

Figure 10 shows the number of dry spells (at least 5 days of rainfall of less than $0.01 mmhr^{-1}$) between 1993 and 2002 for each 0.5° grid cell in December, January, February (DJF). DJF is the dominant wet season over the region and therefore the major growing period for rainfed agriculture. Dry periods within the wet season are important for plant survival and growth. Thus the figure shows the areas where the wet season is prone to interruption. It should be noted that with this definition of dry spells, regions where there are possibly long dry spells without interruption, such as the Namibian Desert show low numbers of dry spells.

5. Summary

A high-resolution 0.1° daily rainfall dataset has been created over southern Africa for the years 1993-2002. This dataset may be used in climate and weather studies where high spatial resolution is important or where a statistical approach requires the use of daily data. A comparison with ground based rainfall measurements (GTS) indicates that the MIRA dataset compares more favourably with GTS measurements than the normalised GPI rainfall estimates in its ability to delineate rain from no rain. However no significant improvement is noted in the ability of the algorithm to distinguish rain rate, compared to normalized GPI. A number of examples of the applicability of the dataset were shown. The Southern Africa data set is available on CD-ROM at http://ltpwww.gsfc.nasa.gov/s2k/html_pages/groups/precip.html.

Acknowledgements.

The authors would like to thank the TAMSAT group at the University of Reading for the provision of Meteosat data. Support for R. Layberry has been provided by EU funded Water and ecosystem resources in regional development (WERRD) project.

References

Andersson, L., T. Gumbricht, D. Hughes, D.R. Kniveton, S. Ringrose, H. Savenije, M. Todd, J. Wilk, and P. Wolski, 2003: Water Flow Dynamics in the Okavango River Basin and Delta - a Prerequisite for the Ecosystems of the Delta. *Physics and Chemistry of the Earth,* 28, 1165-1172.

Adler, R.F., A.J. Negri, P.R. Keehn, and I.M. Hakkarinen, 1993: Estimation of Monthly Rainfall over Japan and Surrounding Waters from a Comination of Low-Orbit Microwave and Geosynchronous IR Data. *J. Appl. Meteorol.*, 32, 335-356.

Adler R.F., C. Kidd, G. Petty, M. Morissey, and H.M. Goodman, 2001: Intercomparison of global precipitation products: The third Precipitation Intercomparison Project (PIP-3). *Bull. Am. Met. Soc*., 82 (7): 1377-1396.

Arkin, P.A. and B.N. Meisner, 1987: The Relationship between Large Scale Convective Rainfall and Cold Cloud over the Western Hemisphere during 1982-84. *Monthly Weather Review*. Vol.115, 51-74.

Barrett, E.C., R.F. Alder, P. Bauer, W. Berg, A. Chang, R. Ferraro, J. Ferriday, S. Goodman, Y. Hong, J. Janowiak, C. Kidd, , D.R. Kniveton, M. Morrissey, W. Olson, G. Petty, B. Rudolf, A. Shibata, E.A. Smith, and R. Spencer, 1994: The first WetNet

Precipitation Intercomparison Project: Interpretation of results. *Remote Sensing Reviews*, 11(1-4), p303-373.

Bellerby, T.J., M.C. Todd, D.R. Kniveton, C. Kidd, 2000: Rainfall estimation from a combination of TRMM Precipitation Radar and GOES Multi-Spectral Imagery through the use of an artificial neural network. *J. Appl. Meteorol.*, *39*, 12, 2115-2128.

Ciach G.J., E. Habib, W.F. Krajewski, 2003: Zero-covariance hypothesis in the error variance separation method of radar rainfall verification. Advances in Water Resources, 26 (5): 573-580.

Camberlin, P., S. Janicot, and I. Poccard, 2001: Seasonality and atmospheric dynamics of the teleconnection between African rainfall and tropical sea surface temperature: Atlantic vs. ENSO, *Int. J. Climatol.*, 21, 973-1005.

Ebert E.E., M.J. Manton, P.A. Arkin, R.J. Allam, G.E. Holpin, and A. Gruber, 1996: Results from the GPCP algorithm intercomparison programme. *Bull. Am. Met. Soc.*, 77 (12): 2875-2887.

Ferraro R, and Q.H. Li, 2002: Detailed analysis of the error associated with the rainfall retrieved by the NOAA/NESDIS Special Sensor Microwave/Imager algorithm 2. Rainfall over land. *J. Geophys. Res.-Atmos.*, 107 (D23): art. no. 4680.

Gebremichael M., W.F. Krajewski, M. Morrissey, D. Langerud, G.J. Huffman, R. Adler, 2003: Error uncertainty analysis of GPCP monthly rainfall products: A data-based simulation study. *Journal of Applied Meteorology*, 42 (12): 1837-1848.

Houghton, J.T., Y. Ding, D.J. Griggs, M. Noguer, P.J. van der Linden, and D. Xiaosu, (Eds.) 2001: *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change* (IPCC). Cambridge University Press, Cambridge, 944p.

Huffman GJ, R.F. Adler, P. Arkin, A. Chang, R. Ferraro, A. Gruber, J. Janowiak, A. McNab, B. Rudolf, U. Schneider, 1997: The Global Precipitation Climatology Project (GPCP) Combined Precipitation Dataset. *B. Am. Meteorol. Soc*., 78 (1), 5-20.

Joyce R.J., J.E. Janowiak, P.A. Arkin, and P.P. Xie, 2004: CMORPH: A method that produces global precipitation estimates from passive microwave and infrared data at high spatial and temporal resolution. *J. Hydrometeorology*, 5 (3): 487-503.

Lebel T, and A. Amani, 1999: Rainfall estimation in the Sahel: What is the ground truth? *J. Appl. Meteorol*., 38(5), 555-568.

Kidd, C., 2001: Satellite rainfall climatology: A review. *Int. J. Climatol.*, 21 (9), 1041-1066.

Kummerow C., and L. Giglio, 1994: A passive microwave technique for estimating rainfall and vertical structure information from space. 1. Algorithm Description. *J. Appl. Met*. 33(1): 3-18.

Kummerow C., W.S. Olson, and L. Giglio, 1996: A simplified scheme for obtaining precipitation and vertical hydrometeor profiles from passive microwave sensors. *IEEE Trans. Geosci. Rem. Sens.*, 34(5): 1213-1232.

Kummerow C., Y. Hong, W.S. Olson, S. Yang, R.F. Adler, J. McCollum, R. Ferraro, G. Petty, D.B. Shin, and T.T. Wilheit, 2001: The evolution of the Goddard profiling algorithm (GPROF) for rainfall estimation from passive microwave sensors. *J. Appl. Met*., 40 (11): 1801-1820

Mason, S.J., and M.R. Jury, 1997: Climatic variability and change over southern Africa: A reflection on underlying processes. *Progress in Physical Geography***,** 21 (1), 23-50.

Miller S.W., P.A. Arkin, and R. Joyce, 2001: A combined microwave/infrared rain rate algorithm. *Int. J. Rem. Sens*., 22 (17): 3285-3307

New, M., M. Hulme, and P. Jones, 2000: Representing twentieth-century space-time climate variability. Part II: Development of 1901-96 monthly grids of terrestrial surface climate. *J. Climate*, 13(13), 2217-2238.

Sorooshian, S., K. Hsu, X. Gao, H.V. Gupta, B. Imam and D. Braithwaite, 2000. Evaluation of PERSIANN system satellite-based estimates of tropical rainfall, *Bulletin of the American Meteorological Society*, 81(9), 2035-2046.

Smith, E.A., J.E. Lamm, R. Adler, J. Alishouse, K. Aonashi, E.C. Barrett, P. Bauer, W. Berg, A. Chang, R. Ferraro, J. Ferriday, S. Goodman, N. Grody, C. Kidd, D.R. Kniveton, C. Kummerow, G. Liu, F. Marzano, A. Mugnai, W. Olson, G. Petty, A. Shibata, R. Spencer, F. Wentz, T. Wilheit, E. Zipser, 1998: Results of WetNet PIP-2. *Journal of Atmospheric Sciences*, 55(9), 1483-1536.

Stanski, H.R., L.J. Wilson, and W.R. Burrows, 1989: *Survey of common verification methods in meteorology*. World Weather Watch Tech. Rept. No.8, WMO/TD No.358, WMO, Geneva, 114 pp.

Todd, M.C., C. Kidd, D.R. Kniveton, T.J. Bellerby, 2001: A combined satellite infrared and passive microwave technique for estimation of small scale rainfall. *J. Atmos. Ocean. Technol*., 18 (5), 742-755.

Xu, L., X. Gao, S. Sorooshian, P.A. Arkin, and B. Imam, 1999: A Microwave Infrared Threshold Techique to Improve the GOES Precipitation Index. *J. Appl. Meteorol.*. 38, 569-579.

Tables

Table 1. Results of contingency table analysis for MIRA and normalised GPI with GTS for the year 2000.

Table 2. HSS results for MIRA and normalised GPI with GTS against number of GTS stations used in the comparison.

Figures

Figure 1. Spatial coverage of the GTS gauge dataset (1990-2000). 0.5 degree cells containing one or more gauges are marked.

Figure 2. Tb-rain rate relationship for January 1993 for a grid cell located in South Africa.

Figure 3. Threshold temperature (mean, K) for December-January-February.

Figure 4. MIRA daily POD (2000)

Figure 5a. Mean rainfall over Southern Africa for January 1999 from MIRA (mm hr^{-1}).

Figure 5b. Mean rainfall over Southern Africa for January 1999 from normalised GPI (mm hr^{-1}).

Figure 5c. Mean rainfall over Southern Africa for January 1999 from GTS (mm hr^{-1}).

Figure 6. Scatterplot of MIRA vs. rain gauge estimates of daily rainfall at 0.5° spatial resolution for the year 2000, for grid cells where there is at least one rain gauge present.

Figure 7. MIRA estimated mean rainfall between 21st and 25th February 2000 - a high rainfall event (mm hr^{-1}).

Figure 8. Coefficient of variation of the daily rainfall over southern Africa for the period 1993-2002.

Figure 9. Difference between the probability of a rain-rain day in an El-Nino year minus the same probability in a La-Nina year for the period 1993-2002.

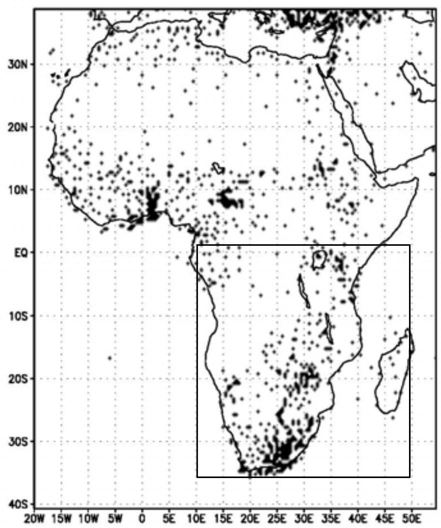
Figure 10. The number of dry spells (at least 5 days) between 1993 and 2002 for each 0.5 degree grid cell in the wet season (December–January-February).

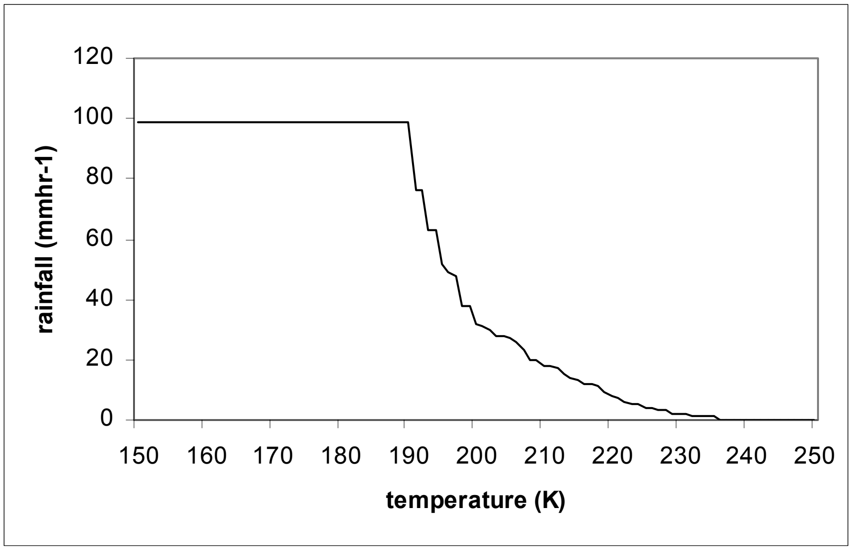
Table 1 Results of contingency table analysis for MIRA and normalised GPI with GTS for the year 2000

	MIRA-GTS	GPI-GTS	
Accuracy	0.75	0.71	range 0 - 1, perfect score 1
Bias	1.14	0.85	range 0 - Inf, perfect score 1
POD	0.71	0.60	range 0 - 1, perfect score 1
FAR	0.38	0.29	range 0 - 1, perfect score 0
CSI	0.50	0.48	range 0 - 1, perfect score 1, 0 indicates no skill
ETS	0.31	0.25	range -1/3 - 1, perfect score 1, 0 indicates no skill
HK	0.49	0.40	range -1 - 1, perfect score 1, 0 indicates no skill
HSS	0.47	0.40	range -Inf - 1, perfect score 1, 0 indicates no skill
OdR	8.46	5.91	range 0 - Inf, perfect score Inf, 1 indicates no skill

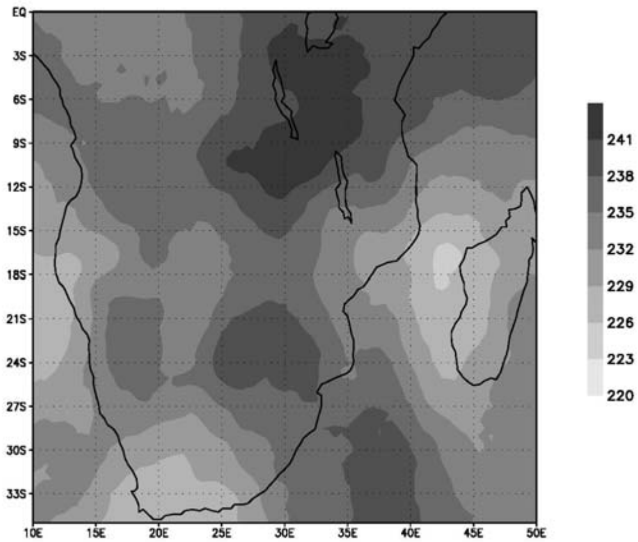
Table 2 HSS results for MIRA and normalised GPI with GTS against number of GTS stations used in the comparison

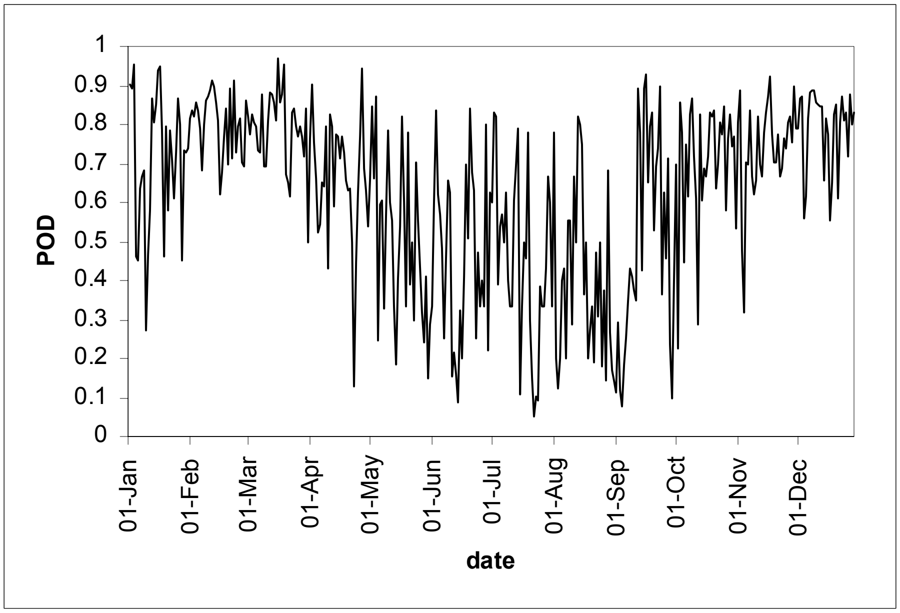
	HSS MIRA-GTS	HSS GPI-GTS	Gauges used (daily mean)
1993	0.46	0.34	187
1994	0.48	0.34	192
1995	0.34	0.30	153
1996	0.40	0.30	170
1997	0.42	0.40	199
1998	0.47	0.37	252
1999	0.46	0.35	223
2000	0.47	0.40	227

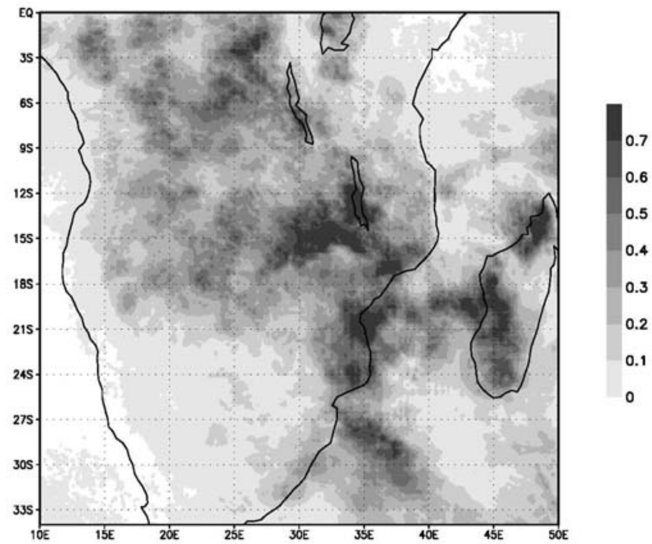

30N
20N
10N
EQ
10S
20S
30S
40S
20W 15W 10W 5W 0 5E 10E 15E 20E 25E 30E 35E 40E 45E 50E

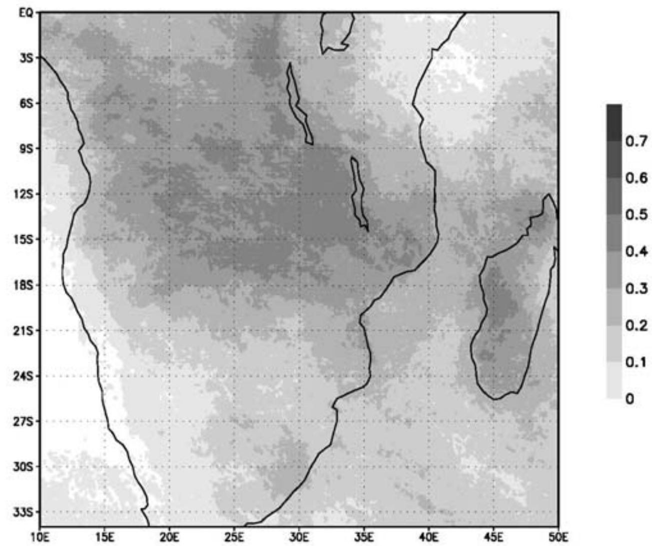


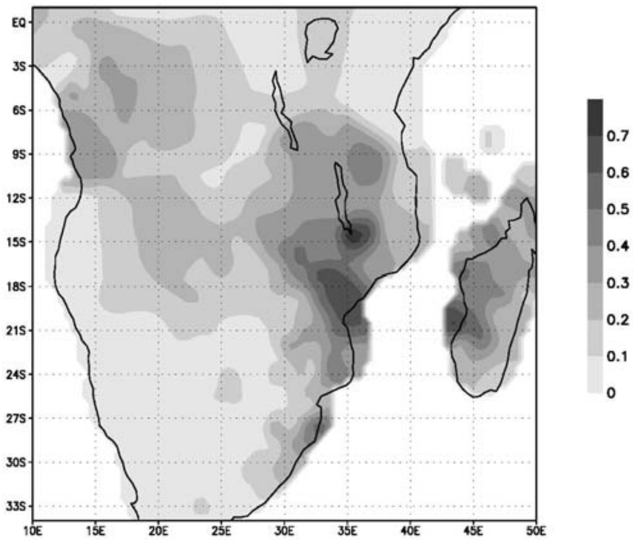
120
100
80
60
40
20
0
rainfall (mmhr-1)
150 160 170 180 190 200 210 220 230 240 250
temperature (K)

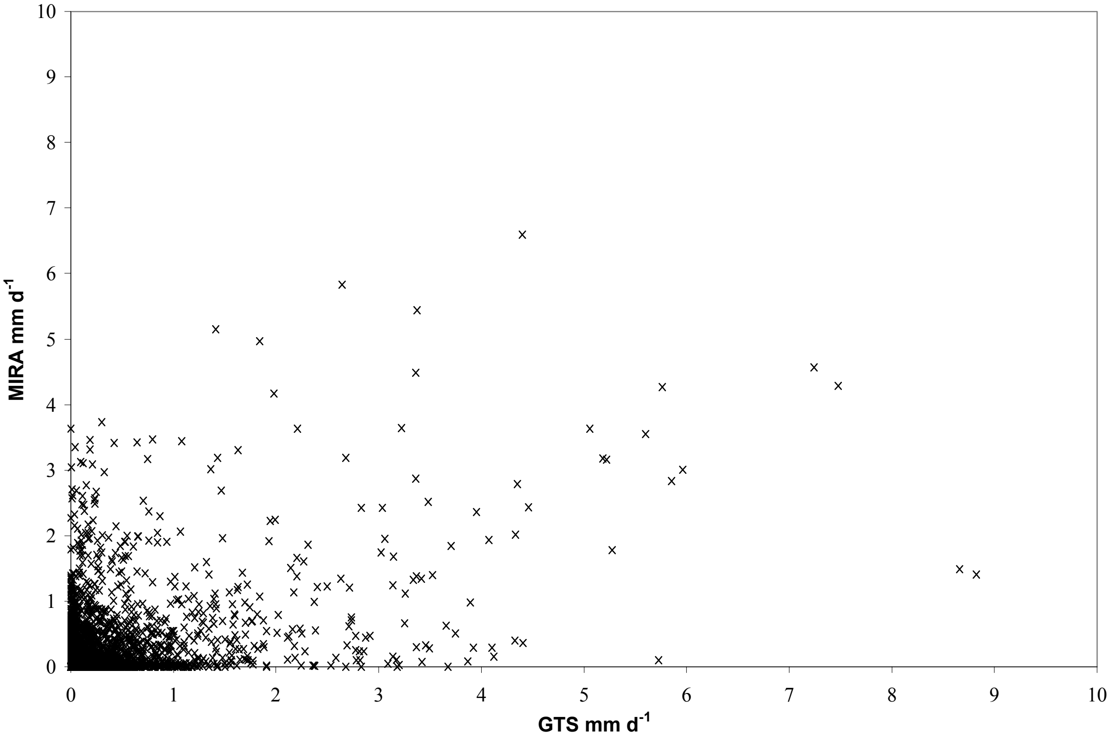

EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
241
238
235
232
229
226
223
220

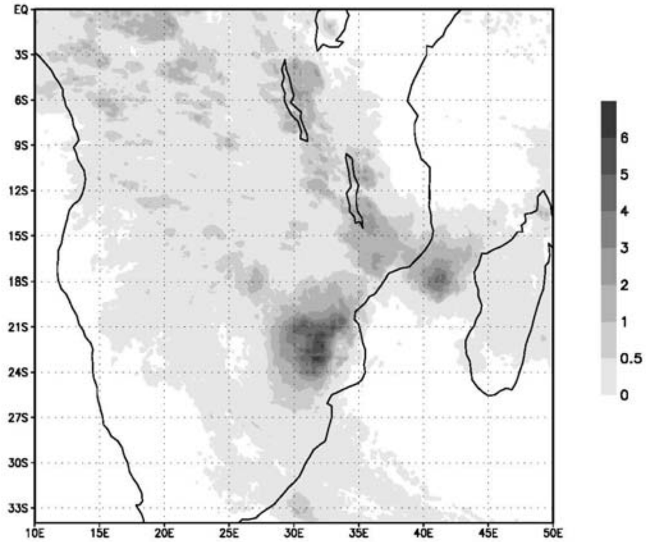

POD
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
01-Jan
01-Feb
01-Mar
01-Apr
01-May
01-Jun
01-Jul
01-Aug
01-Sep
01-Oct
01-Nov
01-Dec
date

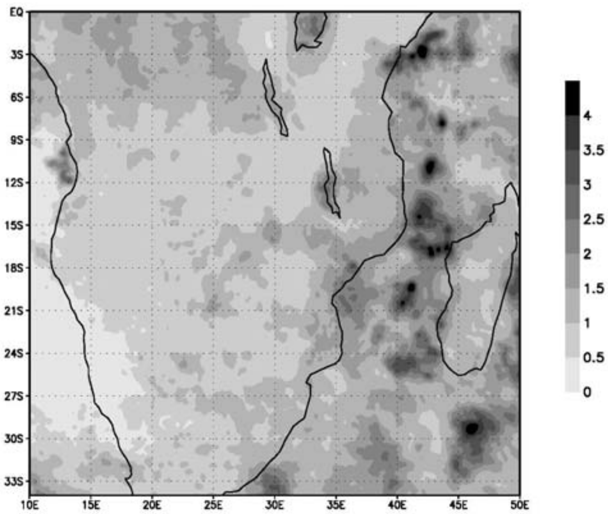

EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

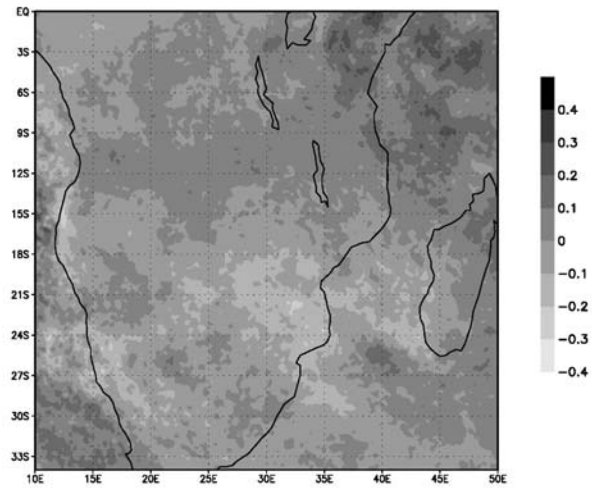

EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0

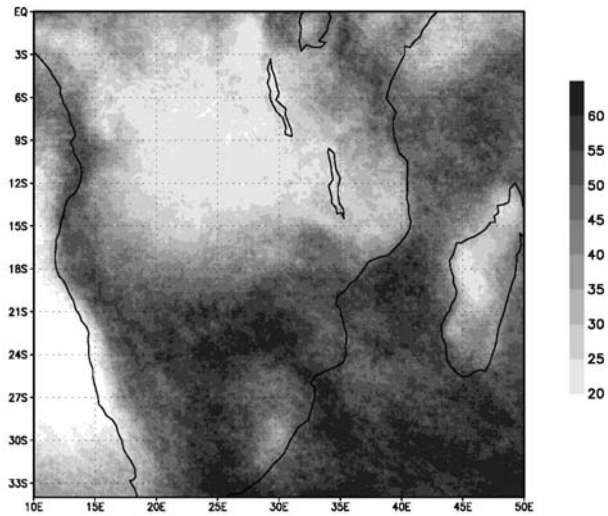

EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0


MIRA mm d-1
GTS mm d-1
0
1
2
3
4
5
6
7
8
9
10


EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
6
5
4
3
2
1
0.5
0


EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
4
3.5
3
2.5
2
1.5
1
0.5
0


EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
0.4
0.3
0.2
0.1
0
−0.1
−0.2
−0.3
−0.4


EQ
3S
6S
9S
12S
15S
18S
21S
24S
27S
30S
33S
10E
15E
20E
25E
30E
35E
40E
45E
50E
60
55
50
45
40
35
30
25
20